We Claim:

1. A EGFR or EGFRKD protein, or a functional EGFRKD protein subunit, in crystalline form.

2. The crystalline protein or functional protein subunit of claim 1, which is a heavy-atom derivative crystal.

3. The crystalline protein or functional protein subunit of claim 2, in which EGFRKD protein is a mutant.

4. The crystalline protein of claim 3, which is characterized by a set of structural coordinates that is substantially similar to the set of structural coordinates of Fig. 4 or Fig. 5.

5. A crystal comprising EGFR protein and a ligand.

6. A method of identifying a ligand that binds EGFR protein, comprising;

   a) forming a co-crystal of a test ligand and EGFR protein;

   b) analyzing said co-crystal using X-ray crystallography; and

   c) using said analysis to determine whether said test ligand binds EGFR protein.

7. The method of claim 6 wherein said co-crystal is obtained by soaking a EGFR protein crystal in a solution comprising said test ligand.

8. The method of claim 7 wherein said co-crystal is obtained by co-crystallizing EGFR protein in the presence of said test ligand.

9. A machine-readable medium embedded with information that corresponds to a three-dimensional structural representation of a crystalline protein of claim 1.

10. The machine-readable medium of claim 9, embedded with the molecular structural coordinates of Fig. 4 or Fig. 5, or at least 50% of the coordinates thereof.

11. The machine-readable medium of claim 9, embedded with the molecular structural coordinates of Fig. 4 or Fig. 5, or at least 80% of the coordinates thereof.

12. The machine-readable medium of claim 9, embedded with the molecular structural coordinates of a protein molecule comprising a EGFRKD protein binding pocket, wherein said binding pocket comprises at least three amino acids selected from the group consisting of Leu24, Val32, Ala49, Thr96, Gln97, Met99, Leu150, Thr160, Gly25, Ser26, Gly27, Ala28, Phe29, Gly30, Lys34, Lys51, Leu98, Gly102, Asn148, Asp161, Leu164, Cys81, Leu94, Pro100, Cys103, Asp143, and Arg147, having the structural coordinates of Fig. 4 or Fig. 5, or by the structural coordinates of a binding pocket homolog, wherein said the root mean square deviation of the backbone atoms of the amino acid residues of said binding pocket and said binding pocket homolog is less than 2.0Å.

13. The machine-readable medium of claim 12, wherein said binding pocket comprises Leu24, Val32, Ala49, Thr96, Gln97, Met99, Leu150, and Thr160, according to the sequence of Fig. 4 or Fig. 5.

14. The machine-readable medium of claim 13, wherein said binding pocket further comprises Gly25, Ser26, Gly27, Ala28, Phe29, Gly30, Lys34, Lys51, Leu98, Gly102, Asn148, Asp161, and Leu164 according to the sequence of Fig. 4 or Fig. 5.

15. The machine-readable medium of claim 14, wherein said binding pocket further comprises Cys81, Leu94, Pro100, Cys103, Asp143, and Arg147 according to the sequence of Fig. 4 or Fig. 5.

16 A method of producing a computer readable database comprising the three-dimensional molecular structural coordinates of a binding pocket of a EGFRKD protein, said method comprising

a) obtaining three-dimensional structural coordinates defining said protein or a binding pocket of said protein, from a crystal of said protein; and

b) introducing said structural coordinates into a computer to produce a database containing the molecular structural coordinates of said protein or said binding pocket.

17. A computer readable database produced by claim 16.

18. A method of producing a computer readable database comprising a representation of a compound capable of binding a binding pocket of a EGFRKD protein, said method comprising

a) introducing into a computer program a computer readable database produced by claim 16;

b) generating a three-dimensional representation of a binding pocket of said EGFRKD protein in said computer program;

c) superimposing a three-dimensional model of at least one binding test compound on said representation of the binding pocket;

d) assessing whether said test compound model fits spatially into the binding pocket of said EGFRKD protein; and

e) storing a representation of a compound that fits into the binding pocket into a computer readable database.

19. A method of producing a computer readable database comprising a representation of a binding pocket of a EGFRKD protein in a co-crystal with a compound, said method comprising

a) preparing a binding test compound represented in a computer readable database produced by claim 18;

b) forming a co-crystal of said compound with a protein comprising a binding pocket of a EGFRKD protein;

c) obtaining the structural coordinates of said binding pocket in said co-crystal; and

d) introducing the structural coordinates of said binding pocket or said co-crystal into a computer-readable database.

20. A computer readable database produced by claim 18.

21. A method of modulating EGFRKD protein activity comprising contacting said EGFRKD with a compound, wherein said compound is represented in a database produced by the method of claim 18.

22. A method of producing a compound comprising a three-dimensional molecular structure represented by the coordinates contained in a computer readable database produced by claim 18 comprising synthesizing said compound wherein said compound binds in a binding pocket of EGFRKD protein.

23. A method of modulating EGFRKD protein activity, comprising contacting said EGFRKD protein with a compound produced by claim 22.

24. A method of identifying an activator or inhibitor of a protein that comprises a EGFRKD active site or binding pocket, comprising

a) producing a compound according to claim 22;

b) contacting said compound with a protein that comprises a EGFRKD active site or binding pocket; and

c) determining whether the potential modulator activates or inhibits the activity of said protein.

25. A method for homology modeling the structure of a EGFRKD protein homolog comprising:

a) aligning the amino acid sequence of a EGFRKD protein homolog with an amino acid sequence of EGFRKD protein;

b) incorporating the sequence of the EGFRKD protein homolog into a model of the structure of EGFRKD protein, wherein said model has the same structural coordinates as the structural coordinates of a crystalline protein of claim 1, or the

structural coordinates of Fig. 4 or Fig. 5, or wherein the structural coordinates of said model's alpha-carbon atoms have a root mean square deviation from the structural coordinates of Fig. 4 or Fig. 5, of less than 2.0Å to yield a preliminary model of said homolog;

c) subjecting the preliminary model to energy minimization to yield an energy minimized model; and

d) remodeling regions of the energy minimized model where stereochemistry restraints are violated to yield a final model of said homolog.

26. A method for identifying a compound that binds EGFRKD protein comprising:

a) providing a computer modeling program with a set of structural coordinates or a three dimensional conformation for a molecule that comprises a binding pocket of a crystalline protein of claim 1, or a homolog thereof;

b) providing a said computer modeling program with a set of structural coordinates of a chemical entity;

c) using said computer modeling program to evaluate the potential binding or interfering interactions between the chemical entity and said binding pocket; and

d) determining whether said chemical entity potentially binds to or interferes with said protein or homolog.

27. A method for designing a compound that binds EGFRKD protein comprising:

a) providing a computer modeling program with a set of structural coordinates, or a three dimensional conformation derived therefrom, for a molecule that comprises a binding pocket comprising the structural coordinates of a binding pocket of a crystalline protein of claim 1, or a homolog thereof;

b) computationally building a chemical entity represented by set of structural coordinates; and

c) determining whether the chemical entity is expected to bind to said molecule.

28. The method of claim 27, wherein determining whether the chemical entity potentially binds to said molecule comprises performing a fitting operation between the chemical entity and a binding pocket of the molecule; and

computationally analyzing the results of the fitting operation to quantify the association between the chemical entity and the binding pocket.

29. A method of producing a mutant EGFRKD protein, having an altered property relative to EGFRKD protein, comprising,

a) constructing a three-dimensional structure of EGFRKD protein having structural coordinates selected from the group consisting of the structural coordinates of a crystalline protein of claim 1, the structural coordinates of Fig. 4 or Fig. 5, and the structural coordinates of a protein having a root mean square deviation of the alpha carbon atoms of said protein of less than 2.0Å when compared to the structural coordinates of Fig. 4 or Fig. 5;

b) using modeling methods to identify in the three-dimensional structure at least one structural part of the EGFRKD protein molecule wherein an alteration in said structural part is predicted to result in said altered property;

c) providing a nucleic acid molecule coding for a EGFRKD mutant protein having a modified sequence that encodes a deletion, insertion, or substitution of one or more amino acids at a position corresponding to said structural part; and

d) expressing said nucleic acid molecule to produce said mutant;

wherein said mutant has at least one altered property relative to the parent.

30. A method of producing a computer readable database containing the three-dimensional molecular structural coordinates of a compound capable of binding the active site or binding pocket of a protein molecule, said method comprising

a) introducing into a computer program a computer readable database produced by claim 16;

b) generating a three-dimensional representation of the active site or binding pocket of said EGFRKD protein in said computer program;

c) superimposing a three-dimensional model of at least one binding test compound on said representation of the active site or binding pocket;

d) assessing whether said test compound model fits spatially into the active site or binding pocket of said EGFRKD protein;

e) assessing whether a compound that fits will fit a three-dimensional model of another protein, the structural coordinates of which are also introduced into said computer program and used to generate a three-dimensional representation of the other protein; and

f) storing the three-dimensional molecular structural coordinates of a model that does not fit the other protein into a computer readable database.

31. A method for determining whether a compound binds EGFRKD protein, comprising,

a) providing a computer modeling program with a set of structural coordinates or a three dimensional conformation for a molecule that comprises a binding pocket of a crystalline protein of claim 1, EGFRKD protein, or a homolog thereof;

b) providing a said computer modeling program with a set of structural coordinates of a chemical entity;

c) using said computer modeling program to evaluate the potential binding or interfering interactions between the chemical entity and said binding pocket; and

d) determining whether said chemical entity potentially binds to or interferes with said protein or homolog.

32. A method of producing a computer readable database comprising a representation of a compound capable of binding a binding pocket of a EGFRKD protein, said method comprising

a) introducing into a computer program a computer readable database produced by claim 16;

b) determining a chemical moiety that interacts with said binding pocket;

c) computationally screening a plurality of compounds to determine which compound(s)comprise said moiety as a substructure of said compound(s); and

d) storing a representation of said compound(s) that comprise said substructure into a computer readable database.

33. Crystallizable EGFR protein.

34. A method of purifying EGFR protein linked to a histidine tag comprising:

a) obtaining a translation vector comprising a coding sequence for EGFR protein, linked to a histidine tag;

b) performing size exclusion chromatography; and

c) performing nickel chelating column chromatography.

35. Purified EGFRKD polypeptide.

36. The method of claim 35 wherein said polypeptide is 98% pure.

37. The method of claim 35 wherein said polypeptide is unphosphorylated.

38. A method of purifying EGFR polypeptide, comprising

expressing EGFR in insect cells;

obtaining a soluble protein fraction from said insect cells;

using a two column chromatograph procedure to obtain purified EGFR.

39. An insect cell capable of expressing EGFR.

40. The insect cell of claim 39, wherein said insect cell comprises a vector, wherein said vector comprises a nucleic acid sequence coding for EGFR.